Weekly Report

2018.0619-2018.0624

1.This Week

Power Grid Deep Learning Paper

1.We receive the model result which seems to be very poor when we try to locate the fault on a specific device. With the model trained on the complete training data, the accuracy cannot even reach 60%, which seems like a 'random guess' result. We're making the following efforts to enhance the accuracy:

- generate more data to be trained (we plan to generate at least 100,000 more instances). The generation speed is now about 5,000 instances/day.
- we keep tunes the model parameters (but we consider the data size is the biggest problem because 10,000 instances is still not enough if the classifier have 3,000 classes)

2.When we increases the accuracy of the model trained on the complete training data, we plan to randomly sample 25% nodes in each instance and repeat this process to generate incomplete training data sets for each instance and use these data sets to train the model to try to get a good performance.

SQC Paper

1. Zongzhuang is realizing the CUSUM algorithm with the real case power grid data (data of 36 nodes in the power grid).

Others

1.review the Chinavis survey paper

2.Read the book machine learning written by Zhou Zhihua and learning basic theories (especially mathematical backgrounds) about machine learning and deep learning. (especially about neural networks, ensemble learning and reinforcement learning)

WaveLines Revision Plan:

changes to be made:

1.make clear the definitions of power gris simulation terms like transient simulation stc..(make a form)

2.clearly define all patterns mentioned in this paper and discuss how existing works distinguish these patterns in the related work.

3.revise the evaluation part: add quantitative evaluation of how accurate wavelines can help to find patterns(effenctiveness), how long will take(efficiency) and complete understanding (comprehensiveness).

4.explain the data preprocessing into a whole continuous process.

5.related work: explain why methods used to visualize pairwise variables mentioned are not used in this paper.(use this to guide the alternative designs discussion.)
6.use less words but add a figure to explain the section waveline design trade-off.
7.explain more carefully why we use a bulb metaphor.

problem:
1.reviewer 3 suggest that expert feedbacks are now informative after the first revision but it is still difficult to judge.
2.reviewer 3 challenges the generability of the approach: we explain that it can be applied to multi-variate time series in a pairwise way, but the reviewer regard it as can be only applied to pairwise variables.

Paper Reading

1.HiPiler: Visual Exploration of Large Genome Interaction Matrices with Interactive Small Multiples

This paper offers convenient and excellent interactions. Following the 'overview first, zoom & filter, details on demand' mantra, this paper includes interactions of selection, browse, rematch, abstract & concretization, filter and correlation. The interaction designs are elegant, concise and efficient. Especially the selection design includes abundant interactions methods like drag & drop, lasso, swipe. This paper teaches me that when we are doing visualization papers, we need to provide abundant and concise interactions, which has always ignored by myself.

2.Recent Research Advances on Interactive Machine Learning

This paper summarizes the recent visual analytics works on interactive machine learning tasks like pattern mining, information retrieval, anomaly detection and topic analysis. The challenges proposed at the end of the paper attracts me: wide adoption of models, improve data quality, explainability for experts and practitioners, monitor training process. Among these tasks, the last three seems to me that there're still a lot of work can be done to make advances to these problems.

3.Relational inductive biases, deep learning, and graph networks

This paper suggests that relational inductive bias exists in most existing deep learning architectures. As a result, the authors propose a deep learning building block: graph network, which can handle more structured representations such as sets or graphs. This block can be integrated with any existing models and I'm thinking it might help to increase the classification accuracy of our model. I'll go through this paper more carefully next week (especially the appendix deduction) to find a way to use this block in our model.

4.Machine Learning (Book written by Zhou Hua)

Read the chapter of neural network and ensemble learning.

2.Progress

Work	Deadline	Progress
Power grid paper with Deeping learning	-	Improving the model accuracy.
SQC Paper	-	About to started.
WaveLine revision	ASAP	Make a revision plan.